Diagnosis of Eye Disease Through Fundus Image Analysis and Machine Learning

Synopsis of the PhD Thesis

Submitted by

Sachin Chavan

(Roll No P006)

Department of Computer Engineering

Supervisor

Dr. Nitin Choubey

submitted in partial fulfilment of the requirements of

DOCTOR OF PHILOSOPHY (Engineering)

MUKESH PATEL SCHOOL OF TECHNOLOGY MANAGEMENT & ENGINEERING
SVKM's NARSEE MONJEE INSTITUTE OF MANAGEMENT STUDIES (Declared as Deemed-to-be University Under Section 3 of the UGC Act, 1956)
V. L. Mehta Road, Vile Parle (West)
MUMBAI –400056

May, 2024

TABLE OF CONTENTS

SECTION NO.	TITLE	PAGE NO.
1	ABSTRACT	1
2	INTRODUCTION	2
	2.1 Diabetic Retinopathy (DR)	2
	2.2 Diabetic Macular Edema (DME)	4
	2.3 Glaucoma	5
3	OBJECTIVES	6
4	LITERATURE REVIEW	7
	4.1 Diabetic Retinopathy (DR) Classification	7
	4.2 Diabetic Macular Edema (DME) Classification	9
	4.3 Glaucoma Classification	11
5	CONTRIBUTION	18
	5.1 Diabetic Retinopathy (DR) Classification (Objectives 1 & 2)	18
	5.1.1. Proposed Frame-wise Severity Scale Classification Model	19
	5.1.2. Results and discussion	20
	5.2 Diabetic Macular Edema (DME) Classification (Objective 3)	24
	5.2.1. Proposed DME classification network	25
	5.2.2. Results and discussion	26
	5.3 Glaucoma Classification (Objective 4)	28
	5.3.1. Proposed Glaucoma Detection and Severity Classification	28
	5.3.2. Results and Discussion	30
6	Conclusion	33
7	Future Scope	35
8	Publications	36
9	References	37

1. ABSTRACT

Keywords: Diabetic Retinopathy (DR), Retinal Fundus Images, Diabetic Macular Edema (DME), Deep Learning, Transfer Learning (TL), Fine Tuning (FT), Classification, Glaucoma, Segmentation.

The increasing prevalence of eye diseases such as Diabetic Retinopathy (DR), Glaucoma, and Diabetic Macular Edema (DME) has led to a growing need for accurate and efficient diagnostic tools. In recent years, Deep Learning (DL) approaches have shown promising results in the automatic detection and classification of these diseases. However, these approaches often face challenges such as high dimensional features, small training datasets, and complex grading systems. To solve such problems, this study proposes three novel DL frameworks for DR, glaucoma, and DME diagnosis and classification. The first of them, the Frame-wise Severity Scale Classification Model (FSSCM), is a DL model built in the framework of transfer learning, using EfficientNet B3 base and fine-tuning it with ResNet 101, to determine the severity of DR based on retinal fundus images. The second framework, the Self-Supervised Category Selective Attention Classifier Network (SSCSAC-Net), leverages self-supervised learning to address the limited availability of labelled data for DME classification. The third framework, GlaucoAttent Net, employs a cascaded U-Net architecture with a novel attention mechanism to effectively segment the Optic Disc (OD) and Optic Cup (OC) to assess glaucoma severity. The proposed framework for DR classification was evaluated on three datasets, which are publicly available i.e. Kaggle, Messidor-2 and Messidor-1. The proposed model for DME classification was evaluated on Messidor-1, Messidor-2 and IDRiD datasets. The Glaucoma detection framework was evaluated on ORIGA, REFUGE and G1020. The results show that the proposed frameworks outperform existing methods in terms of accuracy, precision, recall, F1-score, and kappa coefficient. The FSSCM approach achieved high estimation values of 0.981, 0.985, and 0.983 for precision, recall, and F1-score, respectively, on the three datasets. The GlaucoAttent Net approach achieved outstanding performance with 99.67% accuracy, 99.20% precision, 99.35% F-score and a 99.50% recall. The SSCSAC-Net approach surpassed other techniques with an accuracy of 98.7%, precision of 98.6%, and recall of 98.8%. These novel deep learning-based frameworks offer robust and efficient solutions for the accurate detection and classification of DR, glaucoma, and DME, without the need for large annotated datasets. The proposed approaches have the potential to transform clinical diagnostics and enable early disease detection and intervention.

2. INTRODUCTION

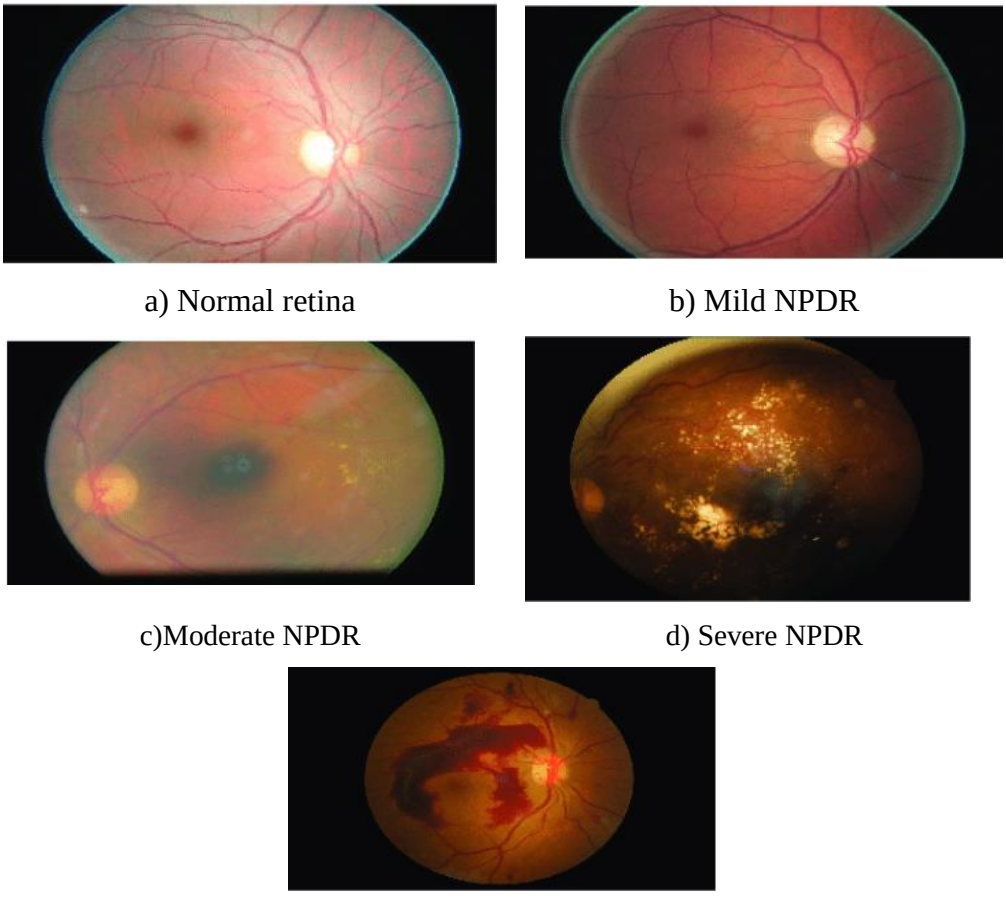
In recent years, it has been observed an exponential increase in diabetic patients. According to World Health Organization, there is an exponential growth in the number of diabetic patients from 1980 to 2014. At actual there were 463 million peoples living with diabetes according to report by International Diabetes Federation in 2019. The worldwide predominance of diabetes among grown-ups more than 18 years old has ascended from 4.7 percent in 1980 to 8.5 percent in 2014. In India pervasiveness of diabetes among grown-ups has ascended from 4.7% in 1980 to 7.8% in 2014. In India the quantity of individuals with diabetes is 102 million in the middle of 2014. Diabetes is a significant reason for visual deficiency, kidney disappointment, coronary episodes, and stroke. Diabetes is one of the reasons for many eye disorders, for example, Diabetic Retinopathy, Glaucoma, Age-related Macular Degeneration, etc.

2.1 Diabetic Retinopathy (DR)

Diabetic retinopathy is a serious eye disease that can cause vision loss and blindness in people with diabetes. As the number of diabetic patients worldwide continues to increase, the number of people with diabetic retinopathy is expected to grow as well. Diabetic retinopathy is caused by damage to the small blood vessels in the retina, a light-sensitive layer at the back of the eye. It is one of the leading causes of vision impairment and blindness in diabetic patients aged 25-75. [1, 2]. A fundus camera takes images of the internal part of the eye, allowing healthcare professionals to evaluate the health of the macula (a central area responsible for sharp vision), the optic disc (where the optic nerve connects to the retina), and the central and peripheral retina [3]. The macula appears as a darker area in the center of the image, while the optic disc is a white spherical area inside the retina. In healthy individuals, certain structures such as the macula fovea, upper disc area, nerve fiber layer, and major blood vessels are not clearly visible in fundus images.

In the medical community, DR grading and classification is the main assessment step to classify based on severity levels determined by the lesions. Generally, there are five phases of disease-related lesions, i.e. Microaneurysms (MA), Hemorrhages (HE), Neovascularization (NV), and Exudates (EX) in fundus images [4]. MA structure is circular with a blood clot size of 100-120µm. This is the earliest visual symptom of DR that appears as the small red dots in the retina. Retinal HE is another disorder due to DR, retinal vein occlusion, and hypertension, where the flowing blood is leaked from the vascular tree in the retina [5]. In the later stage, the unusual development

of blood vessels in the cornea from the limbus is termed NV. EX are fatty deposits situated mainly in the outer layer of the retina. Based on the lesions identified, DR patients are categorized into five classes: class 0-5 is defined as No DR, Mild Non-Proliferative DR (NPDR), Moderate NPDR, Severe NPDR and Proliferative DR (PDR), respectively [6]. The normal and the different phases of DR with increasing severity grades are represented in the figure. 2.1 [7]. From the figure, the retinal changes in the eye are visible where the ophthalmologists estimate the mentioned level to detect the patient's severity. Early identification of DR at the primary stage (mild or moderate stage) is important. From NPDR to the PDR stage, the retinal lesion features are visible clearly by the growth of new blood vessels. DR is a serious problem once it reaches the advanced stage, and sometimes it leads to severe vision loss [8].



a) Normal retina b) Mild NPDR

c)Moderate NPDR d) Severe NPDR

e) PDR

Figure 2.1 Different stages of DR

Recently, many researchers have focused on developing computer-based tools to help diagnose diabetic retinopathy (DR) early. Early detection is important because treatment can prevent up to

90% of cases. However, detecting DR early is challenging. It involves identifying small, subtle changes in fundus images that indicate disease progression. Some existing techniques use deep learning [9-12] to train computers to recognize these changes. These techniques require a large number of labeled images and a long training time. Additionally, the images often have quality issues like reflections, poor contrast, or missing features. There may also be too few images of rare conditions. Despite these challenges, researchers are motivated to develop better systems for grading DR severity. An effective system needs to accurately classify images with high confidence. This will require advanced image processing techniques.

2.2 Diabetic Macular Edema (DME)

DME is a major adverse outcome of diabetes mellitus and is a leading cause of vision impairment in adults of working age worldwide [13]. It arises due to the leakage of fluid from damaged blood vessels into the macula, an area in the center of the retina responsible for sharp and central vision. This fluid accumulation leads to swelling and thickening of the macula, distorting vision [14]. If not promptly diagnosed and treated, DME can lead to severe vision impairment and even blindness, thereby significantly affecting the individual's quality of life and independence [15]. With the advancement of medical imaging and machine learning technologies, automated DME classification has emerged as a promising solution for early detection and regular monitoring of the disease [16]. In this regard, deep learning algorithms have showcased their ability to draw out relevant and distinctive features from retinal images, thereby enhancing the accuracy of DME classification. However, despite the progress, existing methods for DME classification have several inherent limitations [17]. A common issue is the challenge of handling imbalanced datasets, a typical scenario in medical image analysis where the number of normal images far outweighs those showing signs of DME. This imbalance can lead to skewed learning, where the model becomes biased towards the majority class, diminishing its performance on the underrepresented DME classes [18]. Furthermore, many current models heavily rely on extracting high-level or global features, often overlooking critical local or detailed information. These local features, such as specific structural changes or small lesions, can be highly indicative of DME and are vital for accurate diagnosis [19]. Additionally, traditional methods for DME classification often involve the use of manually defined or handcrafted features. While these features can capture certain patterns related to DME, they may not be optimal for detecting the intricate and complex

patterns associated with the disease [20]. Addressing these issues necessitates the development of a more robust and intelligent model. Such a model should be able to leverage both global and local information from retinal images, handle the problem of class imbalance, and adaptively identify and learn the most discriminative features for DME classification.

2.3 Glaucoma

Glaucoma, a leading cause of irreversible blindness globally, necessitates timely detection for effective intervention [21]. Despite its insidious nature, early diagnosis is crucial to preserve visual function and enhance patients' quality of life [22]. Conventional screening methods, such as intraocular pressure measurements and OD examinations, fall short, leading to missed intervention opportunities.

Researchers address glaucoma challenges with diverse machine learning methods like K-Nearest Neighbors (KNN), Naive Bayes, XGBoost and Deep Belief Networks (DBN). Each method has drawbacks—KNN encounters computational intensity, Naive Bayes oversimplifies, XGBoost operates as a black box and DBNs demand substantial labelled data. Additional methods like Support Vector Machines (SVM) offer effective classification but depend on careful parameter tuning. Random Forest [23] handles high-dimensional data but may incur computational expenses. Transfer learning introduces domain adaptation challenges and Recurrent Neural Networks (RNNs) [24] may struggle with limited glaucoma datasets. Ensemble methods like AdaBoost [25] contribute to enhanced performance but elevate complexity and pose interpretability challenges.

Amid these methods, U-Net [26] shines in glaucoma detection, focusing on segmenting key structures crucial for accurate analysis. Unlike traditional approaches, U-Net employs deep learning, allowing it to learn complex patterns efficiently. Despite its value, U-Net has limitations, such as potential inaccuracies in segmenting low-contrast images, occasional oversight of relevant segmentation regions and sensitivity to factors like blood vessels. This necessitates pre-processing, makes precise boundary detection challenging and restricts classifications to binary outcomes.

3. OBJECTIVES

1. To segment anatomical and pathological retinal structures using various image processing techniques. (Paper-01)
2. To analyze the fundus images using computational intelligence methods to grade severity of Diabetic Retinopathy. (Paper-01)
3. To classify the fundus images into normal, non-clinically significant macular edema (NCSME) and clinically significant macular edema (CSME) suspect. (Paper-02)
4. To classify the fundus images into normal and glaucoma images. (Paper-03)

4. LITERATURE REVIEW

In this section, we review the current research on DR, DME and Glaucoma classification and discuss the challenges and opportunities in this field.

4.1 Diabetic Retinopathy (DR) Classification

To detect diabetic retinopathy and make the model simpler, Zago et al. [27] used a patch-based fully convolutional neural network. This type of network processes smaller patches of images instead of the whole image, making it faster. They applied a hard threshold to the labels, treating anything over zero as a positive finding. The model also used a sample selection strategy that focused on difficult cases, improving its performance.

Shanthi et al. [28] described a CNN architecture that modifies Alexnet to classify DR fundus images. The initial convolutional layer processes patches of images, followed by a pooling layer. Unwanted pixels are removed using max-pooling. The modified architecture resizes images based on width and height, with three color channels. The convolutional block calculates a scalar product with weights, and the ReLU layer performs element-wise operations.

The CR-DRNet system by Wu et al. [29] automatically grades the severity of diabetic retinopathy (DR). It uses images of the eye to make this determination. The system has three main parts: image pre-processing, classification, and aggregation. Pre-processing improves the images. The classification part uses a pre-trained network and attention to determine mild or severe DR. The aggregation part combines the results for the final grade.

Riaz et al. [30] presented deep and densely connected neural networks for retinal DR image analysis. The process involves data cleaning, rearranging, and pre-processing. The diagnosis stage utilizes dense patterns of CNN architecture with three dense blocks containing 12 convolutional layers each. The SoftMax function is used to classify fundus images into five stages of DR.

To differentiate normal and DR classes based on feature selection techniques, Gayathri et al. [31], a hybrid feature extraction method using SURF and BRISK techniques is introduced to differentiate normal and DR classes. The SURF descriptor is used for extracting local features and interest points are selected based on the Hessian matrix. The BRISK descriptor considers octaves

and inter-octaves for scale-space key point selection. Both multi-class and binary classification are achieved for the DR images.

Gayathri et al. [32] proposed a combined approach of Anisotropic Dual-Tree Complex Wavelet Transform (ADTCWT) and Haralick feature extraction models for reliable classification. The method performs binary (DR, no DR) and multi-class (mild, moderate, or severe) classification. The Haralick method captures intensity correlation using second-order feature extraction, while the ADTCWT method extracts wavelet features with shift-invariance and multi-dimensionality.

A deep neural network-moth search optimization model was proposed by Shankar et al. [33] to solve the optimization problem in deep CNN for detecting and classifying DR images. The model involves applying the histogram approach for image segmentation, using the Resnet-V2 model for feature extraction, and applying DNN-MSO classification. A sparse auto-encoder with the de-noising technique is used as an enhanced version of DNN, and the weight parameter is updated based on the fittest moth with the MSO algorithm.

Table. 4.1 presented the overall summary of the recent classification model on different stages in DR using retinal fundus images.

Table-4.1. Summary of the existing DR classification task

Author/year	**Technique used**	**Features**	**Advantages**	**Limitations**	**Dataset used**
Zago et al. [27], 2020	Patch-based Deep CNN	Perform lesion localization based on patches	Training requires low number of images	Detection of misclassified samples is high	DIARETDB1, Messidor
Shanthi et al. [28], 2019	Modified Alexnet	Optimal solution is obtained in classification stage	Complexity is reduced	Accuracy is low	Messidor
Wu et al. [29], 2020	CF-DRNet	Self-gated attention highlights the lesion area	Data imbalance problem get reduced	Error occurs among severity grades	Kaggle, and IDRiD fundus database
Riaz et al. [30], 2020	Deep & densely	Dense connection provides deeper	Loss gets reduced	Performance limited to specific datasets	Messidor-2, and EyePACS dataset

	connected model	supervision between layers			
Gayathri et al. [31], 2020	SURF & BRISK	MR-MR improves feature selection and ranking	classifier speed is high	Efficient DR classification is needed	IDRiD, MESSIDOR, DIARETDB0
Gayathri et al. [32], 2020	Haralick & ADTCWT	First order and second order statistics extracts directional features	Integrated features improve the classification task	Error occurs, and complexity is high	MESSIDOR, KAGGLE and DIARETDB0
Shankar et al. [33], 2020	DNN-MSO	Optimization problem solved by MSO algorithm	Inception techniques improves training speed	Accuracy is low	MESSIDOR

Existing methods for lesion localization based on patches and self-gated attention have some limitations. They suffer from high detection of misclassified samples and errors among severity grades. Dense connection provides deeper supervision between layers, but it leads to limited performance on specific datasets and low accuracy. The limitations of the overall study involve: (1) the DR grading task was primarily based on lesion extraction, and achieving accurate grading results would be difficult. (2) lesions may be small to detect due to high misclassified samples, which leads to the whole region constraint, and remains a challenging task in the segmentation process, (3) poor image quality due to low resolution and low intensity, (4) directly using all training features causes high computation time and redundancy and reduces classification performance, (5) due to the high complexity of processing several features, it is hard to train features in the model. To solve these problems, an improved pre-trained network model is necessary based on the extracted features and estimate to diagnose the grades of DR classification with high accuracy.

4.2 Diabetic Macular Edema (DME) Classification

An enhanced deep learning model using the SENet154 network with a Squeeze-and-Excitation (SE) block was proposed by Wu et al. [34] to classify DME risk more accurately. The model includes picture enhancement, weighted random sampling, and generates heat maps to highlight

potential lesion locations. However, accuracy may be compromised if the area of interest is off-center, and the model's generalizability is limited due to training on a small dataset.

DeepOCT, introduced by Altan [35], is a method for diagnosing retinal diseases. It focuses on removing non-retinal areas from OCT images, applying transfer learning, and using an optimized lightweight CNN framework with shallow dense layers. The method achieves good classification performance and employs image preprocessing techniques such as 3D filtering and flattening of retinal layers. It also includes the visualization of significant regions through feature activation maps. However, the effectiveness of these preprocessing techniques may vary depending on the input images' quality and attributes.

Wu et al. [36] developed MD-ResNet, a deep learning model that uses a polar coordinate system for improved classification accuracy and fast processing. The method includes an enhanced Faster R-CNN network for macular center localization and a multi-scale approach for DME severity estimation. However, in cases where the image contains structures resembling the macular center, the method may inaccurately locate it, leading to erroneous DME grading and potential treatment delays.

Saini et al., [37], developed a Lesion based CNN (LCNN) algorithm for improved prediction and lesion detection, focusing on local information. LCNN uses complex features like size, magnitude, and shape, with an attention map during training for precise lesion region prediction. The attention map enhances feature extraction and ensures accurate lesion information extraction. However, they noted challenges with increasing validation loss as the number of samples grew, which might affect generalization on larger datasets. Table 4.2 provides a comprehensive summary of the existing literature.

Table 4.2. Comprehensive summary of existing literature.

Author	**Method**	**Dataset**	**Disadvantage**
Wu et al. [34], 2022	SENet154	Messidor	The mask used in the approach may obscure important details in parts of the image

Altan [35], 2022	DeepOCT	Zhang Lab and SERI	Depends on the quality and characteristics of the input images.
Wu et al. [36], 2021	MD-ResNet	**Training:** Kaggle MESSIDOR, **Testing--**, hospital dataset, e-ophtha EX, HEI-MED and MESSIDOR	Fail to precisely locate the macular center if there are structures or abnormalities in the image that resemble the macular center.
Saini *et al.*, [37], 2023	LCNN	UCSD, NEH	Complexity is higher.

A comprehensive DME classification method is needed to address challenges such as the need for large training data, intricate annotations, and handling of class imbalance. Existing methods focus on macula edema and hard exudates, but ignore important features like vessel structure. The ideal method should accurately identify all relevant features, manage class imbalance, and optimize feature maps.

4.3 Glaucoma Classification

In the realm of glaucoma detection, diverse methodologies have been explored, ranging from deep learning models and segmentation approaches to innovative hybrid methods. Each avenue contributes unique insights and challenges, as discussed below:

4.4.1 Domain Adaptation Models

Domain adaptation is a pivotal aspect of developing robust and generalized models, especially in the context of glaucoma detection. Researchers have explored various strategies to address the challenges posed by differences in domain distributions across diverse datasets. The GDA model (Glaucoma Domain Adaptation) introduced by Madadi et al. [38] is an effective deep learning framework for aligning domain distributions and facilitating knowledge transfer. However, it faces challenges related to complexity, substantial computational requirements, and sensitivity to variations in image quality. The mixup Domain Adaptation (mixDA) method proposed by Yan et

al. [39], combines domain adaptation and domain mixup techniques to enhance glaucoma detection performance. It exhibits superior results across diverse datasets but requires further exploration for comprehensive validation. The Adaptive Feature Fusion model introduced by Pathan et al. [40] provides a novel approach to domain adaptation in glaucoma detection. It dynamically combines features from both source and target domains, resulting in improved adaptation capabilities and enhanced performance. However, challenges remain in optimizing the adaptive fusion parameters.

4.4.2 Ensemble and Attention-Based Models

Ensemble and attention-based models represent cutting-edge methodologies in the realm of glaucoma detection, where the integration of multiple models or attention mechanisms aims to elevate overall system performance. Resheed et al. [41] introduced an innovative approach for glaucoma detection using ensemble learning techniques. The approach combined predictions from multiple models, resulting in improved performance and reduced overfitting. However, this method posed challenges in terms of increased training complexity and potential computational requirements. The deep ensemble network with attention mechanism proposed by Liu et al. [42] achieves superior accuracy in glaucoma screening using stereo optic nerve head images. However, it has dependencies on image quality and requires specialized expertise and computational resources. Mrard et al. [43] proposed a model that combines attention mechanisms with ensemble learning for improved glaucoma detection. This innovative approach aims to enhance feature selection and boost the overall generalization of the model. While the integration of attention and ensemble learning holds great promise for performance gains, it also introduces additional complexity in terms of parameter tuning for both the attention mechanisms and the ensemble components.

4.4.3 Segmentation-Based Approaches

In the realm of glaucoma detection, a segmentation-based strategy is employed to pinpoint and outline specific anatomical structures within medical images, with a particular focus on retinal fundus images of the eye. The fundamental objective of segmentation is to isolate crucial regions or structures linked to glaucomatous pathology, such as the OD and OC. Haider et al. [44] proposed a shallow network like ESS-Net and FBSS-Net, designed for OD and OC segmentation in retinal fundus images. This network can achieve efficient segmentation without extensive preprocessing,

but may require careful hyperparameter tuning. Its strengths include accurate segmentation and computational efficiency, but it may face challenges with limited training data and potential sensitivity to hyperparameter settings. Neto et al. [45] used a type of neural network called a U-Net to segment the optic disk and cup in eye images. This helped them develop an effective method for glaucoma screening.

4.4.4 Multi-Task Learning

Multi-task learning in the field of glaucoma detection involves training a single neural network to simultaneously perform multiple related tasks. A multi-task neural network introduced by Hervella et al. [46] was capable of concurrently handling glaucoma classification and OD/cup segmentation. Their approach demonstrated superior performance in challenging scenarios, eliminating the need for pre-processing or post-processing. However, it introduced different computational requirements compared to baseline multi-task approaches. On a parallel note, Vinayaki and Kalaiselvi [47] introduced the Adaptive Multi-Task Learning model, offering a novel solution to address challenges in glaucoma classification and segmentation. While the optimization of adaptive weights remains an area for further investigation, the model exhibits promising potential to enhance adaptability and overall performance.

4.4.5 Deep Learning and Classification Models

These models leverage neural networks and classification techniques to analyze retinal images and identify signs of glaucoma. Thanki [48] introduced a system utilizing DNN with machine learning classifiers for glaucoma classification, achieving high accuracy and sensitivity. However, the system may be sensitive to dataset characteristics, posing challenges related to potential dataset dependencies and model complexity. Veena et al. [49] employed deep learning CNN models for OD and OC segmentation, demonstrating high accuracy, but challenges include accurately capturing minute features and high computing requirements. Feng et al.'s [50] Transferable Features Classifier presents a promising avenue for advancing glaucoma classification through effective transfer learning. While challenges in fine-tuning large models exist, the potential of this model to enhance generalization and classification performance underscores its significance in developing more robust and accurate glaucoma diagnostic tools.

4.4.6 Hybrid and Innovative Approaches

Hybrid and innovative approaches in glaucoma detection integrate diverse methodologies to enhance diagnostic accuracy and overcome challenges. Madhumalini and Devi [51] introduced a fractional gravitational search-based hybrid deep neural network (FGSA-HDNN), achieving superior classification accuracy through unique hybrid features. However, challenges include the demand for extensive datasets and scalability considerations. Agbaloo and Zaccheus [52] explored the Invariant Scattering Convolution Network, achieving high detection correctness while facing challenges related to parameter sensitivity and limited interpretability. Zarza et al. [53] devised a three-stage training methodology for glaucoma detection, showcasing high precision, reliability, and resource efficiency, with challenges and limitations remaining unexplored. D'Souza et al. [54] introduced a transfer learning approach, demonstrating improved performance by utilizing pre-trained models but requiring substantial computational resources. Wu et al. [55] proposed a three-stage training system resulting in high precision, reliability, and consistent predictions, albeit with the need for careful hyperparameter tuning. The table 4.3 shows the summary of papers.

Table 4.3: Summary of the related work for Glaucoma classification

Author	**Methodology**	**Advantages**	**Challenges**
Madadi et al. [38]	GDA Model: Deep Learning with low-rank representation and weighted source classifier	Effective domain alignment and knowledge transfer	Complexity, substantial computational requirements, sensitivity to image quality
Yan et al. [39]	mixDA Method: Hybrid approach leveraging domain adaptation and mixup	Mitigating domain gaps, improving generalization	Sensitivity to parameter settings

Pathan et al. [40]	Adaptive Feature Fusion Model: Dynamically combining features from source and target domains	Improved adaptation capabilities, enhanced overall performance	Challenges in careful optimization of adaptive fusion parameters
Resheed et al. [41]	Ensemble Learning: Combining predictions from multiple models	Robust results, reduced overfitting	Increased complexity of model training, potential computational demands
Liu et al. [42]	Ensemble with Attention Mechanism: Deep ensemble network with attention mechanism	Improved accuracy, stability	Dependency on image quality, need for specialized expertise and resources
Mrad et al. [43]	Attention-Ensemble Fusion Model: Integrating attention mechanisms with ensemble learning	Enhanced performance through fusion	Challenges in parameter tuning for attention mechanisms and ensemble components
Haider et al. [44]	ESS-Net and FBSS-Net: Shallow networks for OD and OC segmentation	Accurate segmentation, computational efficiency	Need for extensive training data, potential sensitivity to hyperparameter tuning
Neto et al. [45]	U-Net models with Inception V3 and Inception ResNet V2: For OD and OC segmentation	Accurate segmentation, efficient glaucoma screening	Sensitivity to image quality, reliance on pre-trained models

Hervella et al. [46]	Multi-task Neural Network: Simultaneous glaucoma classification and OD/cup segmentation	Simultaneous classification and segmentation	Different computational requirements, challenges in computational requirements
Vinayaki and Kalaiselvi [47]	Adaptive Multi-Task Learning: Dynamically adjusting task weights	Improved adaptability and overall performance	Challenges in optimizing adaptive weights
Thanki [48]	DNN with ML Classifiers: Utilizing DNN with machine learning classifiers	Superior performance, dual learning	Sensitivity to dataset characteristics, potential dataset dependencies and model complexity
Veena et al. [49]	Deep Learning CNN Models: For OD and OCsegmentation	Efficient and accurate segmentation	Challenges in accurately capturing minute features, high computing requirements
Feng et al. [50]	Transferable Features Classifier: Leveraging transfer learning	Improved generalization, classification performance	Challenges in fine-tuning large models
Madhumalini and Devi [51]	FGSA-HDNN: Fractional gravitational search-based hybrid deep neural network	Enhanced classification accuracy	Need for extensive datasets, scalability considerations
Agbaloo et al. [52]	Invariant Scattering Convolution Network: Direct feature learning	High detection correctness	Sensitivity to parameter settings, limited interpretability

Zarza et al. [53]	Three-Stage Training: Methodology for glaucoma detection using fundus images	High precision, reliability, resource efficiency	Challenges and limitations remain unexplored
D'Souza et al. [54]	Transfer Learning Approach: Utilizing pre-trained models on large datasets	Capture intricate patterns, enhanced model generalization	Need for substantial computational resources for fine-tuning large models
Wu et al. [55]	Three-Stage Training System: Using state-of-the-art network architectures	Reliable and consistent predictions	Need for careful hyperparameter tuning in each training stage

In glaucoma detection, diverse methodologies tackle challenges such as model complexity, parameter sensitivity, and dataset dependencies. Domain adaptation models like Glaucoma Domain Adaptation (GDA) and mixup Domain Adaptation (mixDA) grapple with complexities, computational demands, and image quality sensitivity. Ensemble and attention-based models, including those by Resheed et al., Liu et al., and Mrad et al., face challenges in training complexity and computational demands, coupled with dependencies on image quality. Segmentation-based approaches, exemplified by ESS-Net, FBSS-Net, U-Net models, and Deep Segmentation Fusion, encounter challenges related to hyperparameter tuning, dataset size and sensitivity to image quality. Multi-task learning models address simultaneous tasks with challenges in computational requirements and adaptive weight optimization. Deep learning and classification models face issues like dataset sensitivity, capturing minute features, high computing requirements, and model complexity. Hybrid and innovative approaches, such as FGSA-HDNN, Invariant Scattering Convolution Network, and three-stage training methodologies, confront challenges in extensive datasets, scalability, parameter sensitivity and interpretability, collectively emphasizing the importance of meticulous parameter tuning and further exploration.

5. CONTRIBUTION

5.1 Diabetic Retinopathy (DR) Classification (Objective 1 & 2)

The main contributions of the proposed work for DR classification are summarized as follows,

1. We present a Frame-wise Severity Scale Classification Model (FSSCM) using an improved pre-trained network to classify the severity classes of retinal health and DR images. First, we consider a two-dimensional image (2D) and derive the approximation of the classification model in a frame-wise block using the Transfer Learning (TL) and Fine Tuning (FT) model. In this, TL can exploit rich feature representations to improve performance on the target task for image classification. Then FT addresses the over-fitting problem and enhances previous results with increasing fine data on the fundus image datasets with high classification accuracy.
2. We explore pre-processing step to standardize the image size and enhance retinal image quality from the raw fundus images. Pre-processing is done to standardize the image size and enhance retinal image quality using circular crop and Gaussian blur. In image processing, the Gaussian blur operation is convolved with a Gaussian filter to smooth or blur the images.
3. We employ a segmentation process to express the contour lines of fundus images using the Chan-Vese algorithm. The Chan-Vese algorithm removes the false bright areas around the HE region. The detection of HE is difficult because of the colour, texture, shape, size and contrast variations of MA and HE. The DR image segmentation of HE is obtained based on the minimized energy concept from the complex background. This algorithm uses the level set that expresses the contour lines and performs the segmentation of colour fundus images.
4. We design an FSSCM using Transfer Learning enabled EfficientNet B3 and Fine Tuning enabled ResNet 101 named TL-EN3 and FT-RN 101. TL-EN3 model is applied to capture high-resolution patterns with high accuracy and integrates FT-RN 101 models to maintain a balance between efficiency and accuracy with fewer parameters. Experimental results demonstrate that several qualitative evaluation metrics validate the framework's utility. The proposed FSSCM integrated results achieve a lower misclassification rate and low training loss than the other state-of-art methods.

5.1.1 Proposed Frame-wise Severity Scale Classification Model

In the proposed framework, we proposed a Frame-wise Severity Scale Classification Model (FSSCM) that involves an improved pre-trained network based on TL and FT with EfficientNet B3 and ResNet 101 model to automatically classify the severity scale of retinal health (No DR) and DR images. Consider a two-dimensional image (2D) and derive the approximation of the classification model in a frame-wise block using the TL and FT model. In this, TL can exploit rich feature representations to improve performance on the target task for image classification. Then FT addresses the over-fitting problem and enhances previous results with increasing fine data on the fundus image datasets with high classification accuracy. This model comprises four processing steps: Image pre-processing and augmentation, Segmentation, Feature extraction, and Classification Initially, the pre-processing step is done to standardize image size and enhance retinal image quality using circular crop and Gaussian blur. Next, image augmentation is processed to generate additional data by four operations, such as flipping, rotation, brightness, and contrast adjustment for model training. Secondly, the segmentation process is employed by the Chan-Vese algorithm to constrain the contour line of the image. Then the feature extraction is performed based on the pixel density of the segmented binary image. Next, the extracted features are set as input for severity scale classification. For this process, the proposed model used transfer learning based on EfficientNet B3 and fine-tuned strategy using ResNet 101. During training, a new layer is added only on top of the fixed layer to increase the model size using TL-EN3. Then update only the last layer of the FT-RN 101 model with a smaller number of parameters. In a real-time scenario, the proposed automatic DR system can help the doctor analyze the fundus image using a grading scale, enabling them to promptly choose the patient's next course of action. Emerging medical technology strongly emphasizes lowering the number of visits to eye specialists, lowering the overall cost of care, and increasing the number of patients seen by each physician. The health care provider can accomplish these goals using an improved pre-trained network model. Accordingly, the enhanced pre-trained network-based tool may speed up the screening of this condition and reduce the need for trained human resources for the same probably be a boon for patients. Overall, the automatic DR system benefits would help patients save time, money, and, ultimately, their eyesight. With the right automation (i.e., decision support systems) in place, preventative measures to safeguard vision can be done early and assist lower both the number of DR issues and the risk of blindness. As a result, the screening will be beneficial at any stage of the illness and help prevent

blindness across 90% of patients. The overview of the proposed FSSCM approach is shown in Figure 5.1.

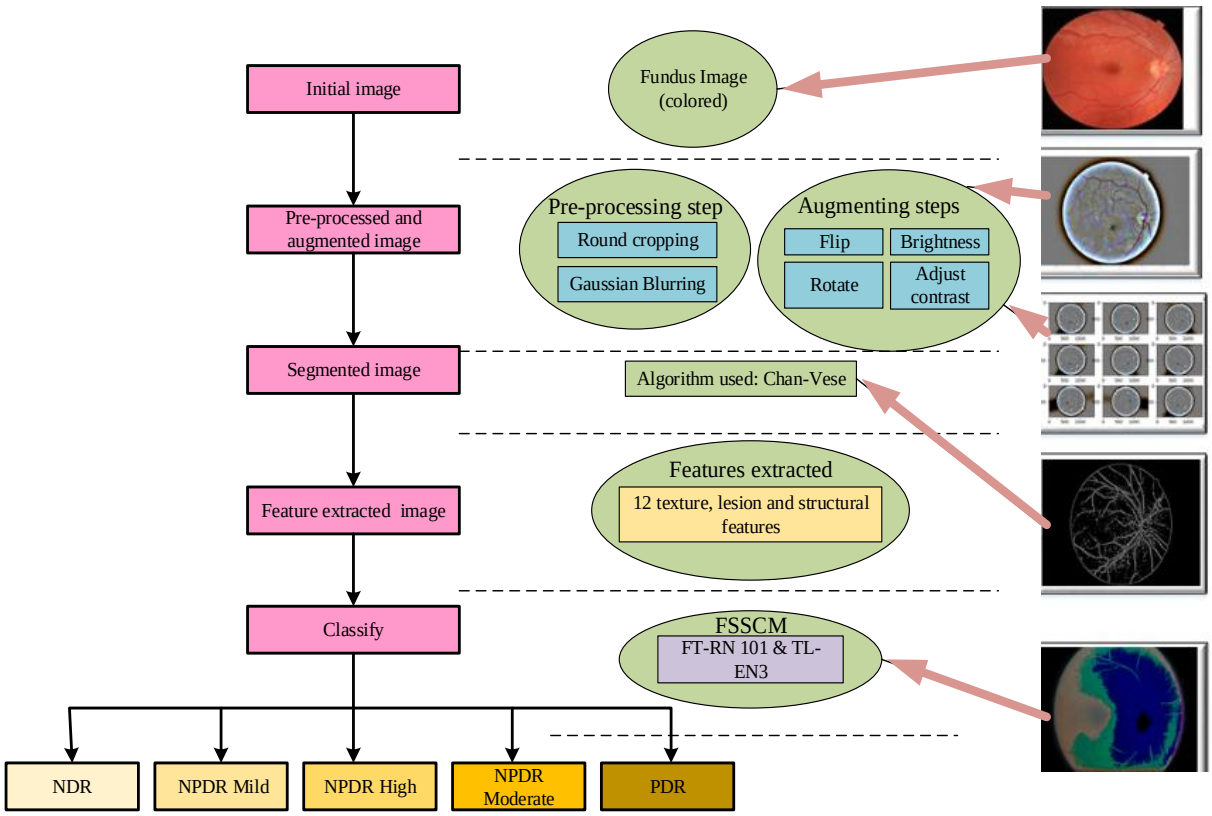


Figure 5.1 Block diagram of the proposed framework

5.1.2 Results and Discussion

The comparative analysis of different classifiers with respect to Kaggle, Messidor-1, and Messidor-2 datasets is mentioned in table 5.1. Several classifiers (RF, SV, J48, MLP, Naïve Bayes and Adaboost are trained independently, each classifier makes a prediction, and all predictions are compared with others. From the analysis, it is clear that the proposed FSSCM model achieves the highest outcome for precision, recall, and f1-score measures while the misclassification rate (FPR) obtains minimum error rate on three datasets.

Table 5.1. Comparison of various efficiency measure with respect to different classifiers

Dataset	**Classifier**	**Precision**	**Recall**	**F1-score**	**FPR**
Kaggle database	RF	0.97	0.97	0.96	0.079
	SVM	0.957	0.962	0.958	**0.049**
	J48	0.97	0.987	0.978	0.25

	FSSCM	**0.981**	**0.985**	**0.983**	0.05
Messidor dataset	RF	0.926	0.918	0.916	0.05
	SVM	0.944	0.942	0.941	0.032
	J48	0.98	0.984	0.98	0.04
	Naïve Bayes	0.948	0.948	0.948	0.051
	Adaboost	0.95	0.945	0.945	0.066
	MLP	0.97	0.979	0.979	**0.021**
	FSSCM	**0.98**	**0.986**	**0.984**	0.08
Messidor-2 dataset	RF	0.92	0.93	0.92	0.24
	SVM	0.89	0.95	0.92	0.26
	J48	0.962	0.97	0.967	**0.019**
	FSSCM	**0.98**	**0.985**	**0.98**	0.07

Table 5.2 shows the details regarding the accuracy and kappa measure for each classifiers using Kaggle, Messidor, and Messidor-2 datasets respectively. The accuracy results obtained for the proposed classifier model is improved than the other classifiers in the DR detection. Also, the performance of kappa measure outperforms in better detection of DR compared to other classifier results. The bold values indicate best results.

Table 5.2 Classification accuracy and kappa coefficient results of each classifier for three datasets

Dataset	**Classifiers**	**Classification accuracy**	***Kappa coefficient***
Kaggle	Random forest	98.01	0.927
	SVM	93.3	0.909
	J48	98.06	0.98
	MLP	94.8	0.99
	FSSCM	**99.02**	**0.993**
Messidor	Random forest	97.8	0.87
	SVM	95.5	0.91
	J48	**99.1**	0.96
	MLP	97.7	0.96
	FSSCM	98.1	**0.97**

Messidor-2	Random forest	95.4	0.95
	SVM	96.1	0.945
	J48	97.4	0.97
	FSSCM	**98.3**	**0.98**

Table 5.3 presents the comparison of system performance evaluated between the proposed and other models conducted with three datasets. The result shows the improved performance than other networks in terms of precision, recall, f1-score, accuracy, and kappa measures. The result clearly indicates that the existing systems achieves lowest performance while the proposed model attains the better classification result of DR grading. The bold values indicate best results and (-) indicates the values are not evaluated.

Table 5.3 Comparison of proposed with earlier approaches

Dataset	**Methods**	**Precision**	**Recall**	**F1-Measure**	**Accuracy**	**Kappa**
Messidor	Modified Alexnet [28]	92.7	0.923	-	96.2	-
	Gayathri et al. [31]	0.98	0.983	0.982	98.2	-
	Gayathri et al. [56]	0.97	0.988	0.979	98.7	0.996
	Gayathri et al. [32]	0.97	0.96	0.96	98.7	-
	M-CNN+J48 [57]	0.99	**0.998**	0.98	99.7	0.996
	FSSCM	**0.997**	0.99	**0.993**	**99.71**	**0.997**
Kaggle	Gayathri et al. [56]	0.97	0.98	0.97	98.9	-
	Gayathri et al. [31]	0.98	0.96	0.97	97.5	-
	Li et al. [58]	-	-	-	86.17	-

	Zeng et al. [59]	-	0.88	-	-	0.82
	M-CNN+J48 [57]	0.98	0.97	0.97	98.7	0.98
	CF-DRNet [29]	0.54	0.91	0.68	83.1	-
	FSSCM	**0.99**	**0.99**	**0.99**	**99.4**	**0.99**
Messidor-2	Lam *et al.* [60]	0.95	-	-	94.4	-
	Sayres *et al.* [61]	0.91	0.94	0.92	95.2	-
	Riaz *et al.* [30]	0.98	0.97	0.97	96.4	-
	FSSCM	**0.98**	**0.99**	**0.98**	**97.8**	**0.98**

This section presents the analysis of the proposed framework for automatic DR severity classification that lies at the core of four processing steps. The proposed Frame-wise Severity Scale Classification Model (FSSCM) involves improved pre-trained networks based on Transfer Learning (TL) and Fine Tuning (FT) with the EfficientNet B3 and ResNet 101 (TL-EN3 and FT-RN101) models for automatic severity classification of DR images, i.e., normal, mild, moderate, severe, and PDR, on three datasets such as the Kaggle, Messidor-1, and Messidor-2 datasets. Initially, the captured images have uneven illumination and an unclear view inside the retinal images. This causes poor image quality. To avoid this problem, the Gaussian blur operation is convolved with a Gaussian filter, which is used to smooth or blur the images. Then an automatic cropping technique is performed to remove the unnecessary area from the image. Furthermore, the colour and position augmentation processes are used to obtain the well-trained deep learning structure and avoid over-fitting issues in the system. After identifying the clear view image, the DR image segmentation of hemorrhages is obtained based on the minimized energy concept from the complex background in assigning disease label to pixels using the Chan-Vese algorithm. After segmentation, the feature extraction process is accomplished to generate unique features of texture, lesion, contrast, structure, and luminance that are obtained from DR and normal retinal images. Finally, the obtained input features are fed as input for classification. For this process, an improved

pre-trained model of TL-EN3 and FT-RN101 is performed to classify the whole features into five classes. In this, the TL-EN3 model is applied to capture high-resolution patterns with high accuracy and integrates FT-RN 101 models to maintain a balance between efficiency and accuracy with fewer parameters. In the step of fine tuning, only the last layer of the model with the smaller number of parameters is updated. This ensures that the output features are learned and reduces the number of parameters from the large dataset. During the training process, a cross-entropy loss function is employed to compute the difference between the input label and the classified result. The performance of the proposed FSSCM approach obtains high estimation values of 0.981 0.985 0.983, 0.98 0.986 0.984, and 0.98 0.985 0.98 in terms of P, R, and F1 on three datasets, respectively. Also, it achieves high estimation results of 99.02 0.993, 98.1 0.97, and 98.3 0.98 in terms of CA and K-score for three datasets, respectively. With a high level of training accuracy and a low level of training loss, the proposed method gets better severity level classification results compared to other models. However, in the classification results, the improved pre-trained network obtains higher accuracy, but it consumes high training time due to the change in the distribution of each layer. So, there is a need to determine the most suitable algorithm for hyper-tuning the parameters in the network model.

5.2 Diabetic Macular Edema (DME) Classification (Objective 3)

Key contributions to the field of DME classification using deep learning:

1. This work introduces the novel SSCSAC-Net, an advanced model that integrates self-supervised learning and category selective attention mechanisms. This innovative architecture aids in precise DME classification.
2. A key contribution of this research is the effective use of unlabeled data for learning robust features related to DME through self-supervised learning. This significantly reduces the dependency on extensive labelled data, thus making the solution more scalable.
3. The work introduces a category-specific attention mechanism that enables the model to concentrate on critical regions of retinal images that are indicative of DME to improve the classification accuracy.
4. The proposed model incorporates a domain-specific layer and an attention mechanism into the ResNet-152 base architecture. This innovation allows the model to better identify and focus on critical DME-related image aspects, thereby enhancing detection and classification accuracy.

5. Finally, the proposed approach is tested on multiple datasets, demonstrating its robustness and generalizability across different data distributions.

Together, these contributions advance the state of the art in automated DME classification, providing a robust and efficient tool that can assist clinicians in diagnosing and treating this serious condition.

5.2.1. Proposed DME classification network

The Self-Supervised Category Selective Attention Classifier Network (SSCSAC-Net) is designed to enhance DME classification accuracy. It includes two integral parts, the Self-supervised orientation aware multi-view discrimination learning network (SSOMDL-Net) and the Category selective attention network (CSA-Net), aiming to address DME classification challenges like need for extensive training data, class imbalance, and feature representation. Utilizing both labelled and unlabeled data, the SSCSAC-Net offers comprehensive DME classification while improving performance and generalization capabilities. Figure 5.2 provides a comprehensive illustration of the SSCSAC-Net's structure.

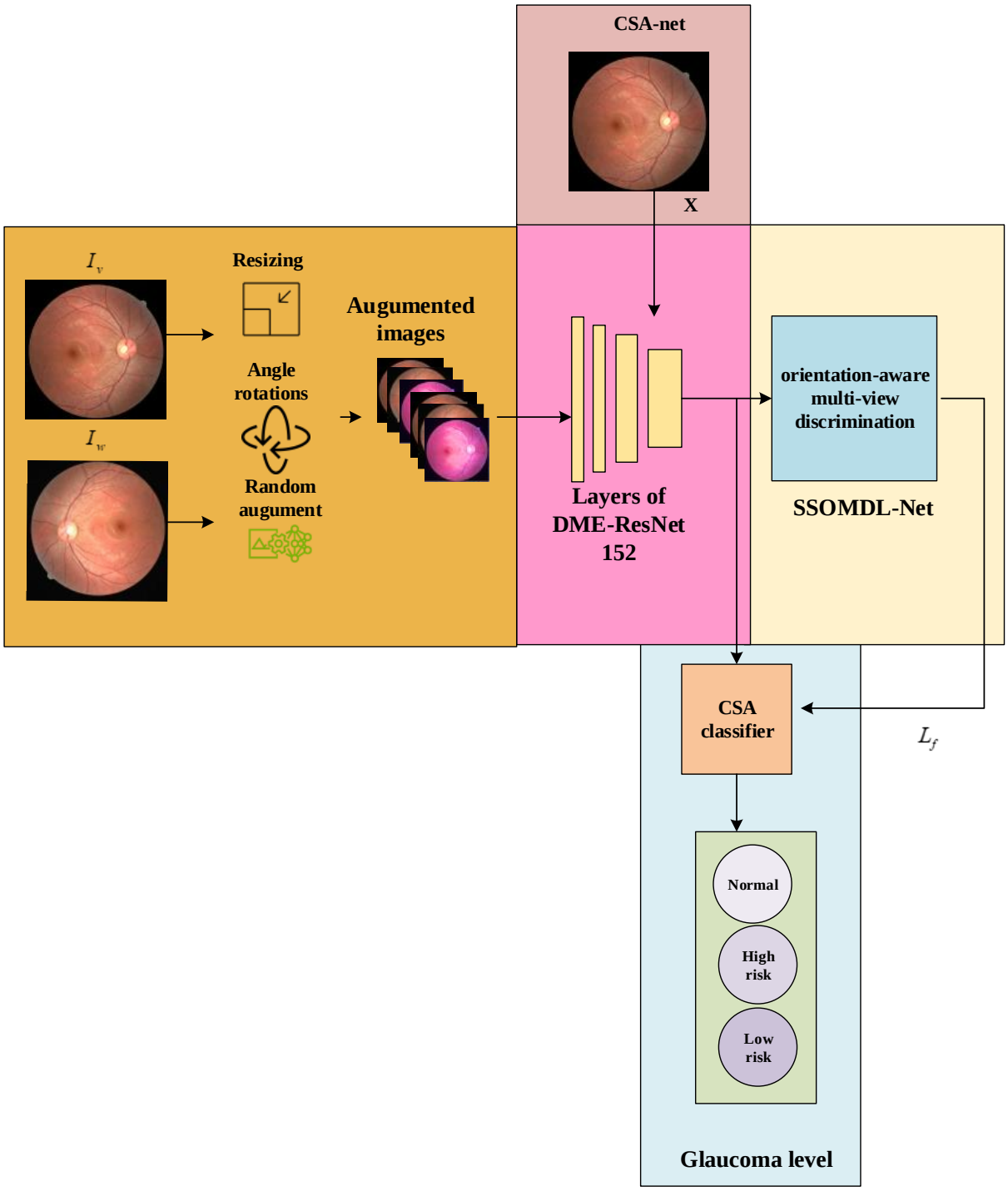


Figure 5.2. Overview of the SSCSAC-Net.

5.2.2 Results and Discussion

The SSCSAC-Net demonstrated robust performance in classifying different severity levels of DME. The SSCSAC-Net's performance was evaluated for each class: normal (class 0), mild/low

risk (class 1), and severe/high risk (class 2). For each category the SSCSAC-Net demonstrated high precision, indicating that it accurately predicted the class labels with minimal false positives. Table 5.4 shows the Class wise performance evaluation of SSCSAC-Net.

Table 5.4 Class wise performance evaluation of SSCSAC-Net.

Evaluation metrics	Precision	Recall	F1-score
Normal (class 0)	99.1	98.7	98.8
Mild/low DME (class 1)	98.3	99.2	98.7
Severe/High DME (class 2)	98.6	98.5	98.5

The precision, recall, and F1-score for class 0 were high, indicating that the model accurately identifies normal cases and correctly differentiates them from DME cases. This is crucial as misclassification of normal cases as DME can lead to unnecessary treatment and anxiety for patients. The model also demonstrated good performance in identifying Class 1 cases. The precision was high, meaning that most cases predicted as mild/low risk by the model were indeed mild/low risk. The recall was also high, showing that the model could correctly identify a significant portion of actual mild/low risk cases. The high F1-score, which is the harmonic mean of precision and recall, indicates a balance between precision and recall for this class. In identifying severe/high risk DME cases (class 2), the model demonstrated exceptional performance. High precision shows that the model accurately predicts severe/high risk cases, while high recall indicates that it captures a significant portion of actual severe/high risk cases in the dataset. The high F1-score suggests a balanced model performance for this class.

Table 5.5 offers a comparison between the proposed method and existing ones across various performance indicators. It is found that the proposed SSCSAC-Net outperforms all the other methods.

Table 5.5 Performance comparison with other models.

Method	Accuracy	Precision	Recall	F-1 score
SCL [62]	84.36	70.51	73.84	-
QEIRV-2 [63]	93.3	-	-	-
Texture attention network [64]	85.1	90.3	92	88.1

Integrated DenseNet with an attention Network [65]	98.06	83.67	-	-
DenseNet169+CBAM+INS [66]	97	97	98.3	-
MSA-Net [67]	98.1	98.3	98.2	-
SENet154 [34]	-	82.2	98.09	-
LCNN [37]	91.2	88.7	87.5	-
Proposed	**98.7**	**98.6**	**98.8**	**98.7**

5.3 Glaucoma Classification (Objective 4)

Proposed glaucoma detection framework which provides the following contributions:

1. The preprocessing step utilizes techniques like NLM denoising, CLAHE and Adaptive Median filtering to optimize fundus images, highlighting features relevant for accurate glaucoma detection.
2. The encoder in GlaucoAttent Net progressively extracts features, starting from low-level details to high-level context and fuses them to create a hierarchical feature map encompassing both granular and global image information.
3. In our proposed GlaucoAttent Net, task-specific knowledge attention prioritizes features essential for precise OD and OC segmentation, guided by tensor knowledge. Spatial attention highlights informative regions in fundus images, while probability maps guide the model by indicating areas more relevant to the segmentation task.
4. We assess our method's performance qualitatively via cross-validation, ROC curve, precision-recall analysis, and GlaucoAttent Net's segmentation performance across datasets. Quantitatively, we undertake an ablation study to assess the influence of individual components.

5.3.1. Proposed Glaucoma Detection and Severity Classification

The proposed Glaucoma detection and severity classification employs a cascaded U-Net with attention mechanisms, named GlaucoAttent Net. Initial steps involve data collection from diverse datasets (ORIGA, REFUGE and G1020) and GFEP using denoising, histogram equalization and adaptive filtering. GlaucoAttent Net integrates encoder, feature fusion and three-stage segmentation paths for accurate OD and OC segmentation. A severity classification module

incorporates derived factors, enabling a comprehensive assessment of glaucoma severity. The resulting output integrates probability maps for OD and OC segmentation, providing a three-stage representation. The severity classification is achieved through fully connected layers and a softmax layer, resulting in a robust glaucoma detection system. An outline of the proposed methodology is illustrated in Figure 5.3.

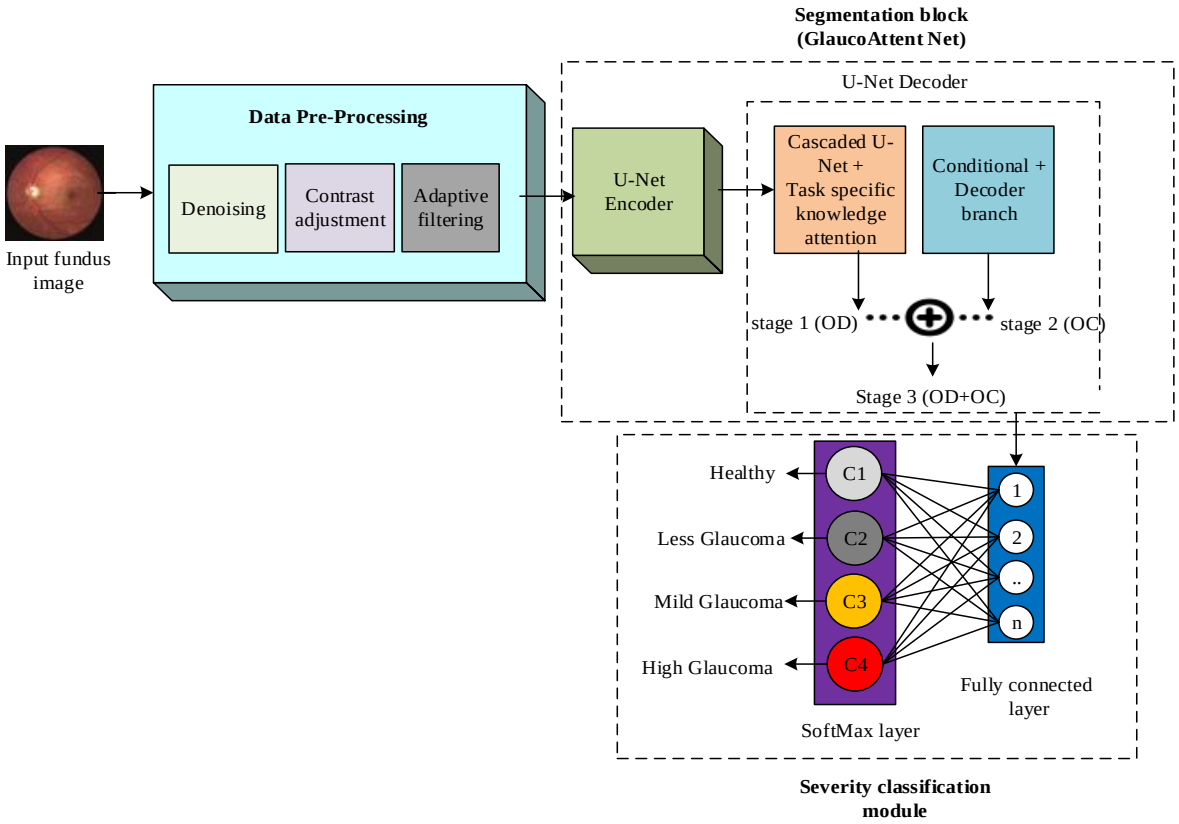


Figure 5.3: Proposed Glaucoma Detection and Severity classification model

5.3.1.1 Data collection

Data collection is a crucial initial step in the proposed glaucoma detection framework, aiming to compile a diverse and representative dataset for training, validation and evaluation. The process involves acquiring fundus images along with corresponding glaucoma annotations from three distinct datasets: ORIGA, G1020 and REFUGE. Each dataset contributes to the creation of a comprehensive repository, ensuring the development and assessment of the glaucoma detection model across various scenarios and populations.

5.3.1.2 Glaucoma Features Enhancement Preprocessing

Glaucoma Features Enhancement Preprocessing optimizes fundus images in the proposed glaucoma detection framework. It employs NLM denoising, CLAHE and Adaptive Median Filtering to enhance glaucomatous features. It contributes to improved accuracy in diagnosing glaucoma by refining image quality and preserving relevant details.

5.3.1.3 GlaucoAttent Net

The grayscale denoised fundus image ($x'(n)$) is the initial input to the encoder path of the U-net based GlaucoAttent Net. The architecture of GlaucoAttent Net is given in figure 5.3. Initially, pixel intensities in the input image are normalized to ensure uniformity and compatibility with the neural network model. The pixel values are scaled linearly to a range between 0 and 1.

a) Encoder Path

In the initial layers of the encoder path, the convolutional operations act as filters to detect local patterns and variations in pixel intensities. The convolutional filters have small receptive fields, enabling them to capture local details in the input image such as, a convolutional filter detect edges by responding to intensity changes in adjacent pixels.

b) Decoder Path

The decoder path in our proposed methodology, GlaucoAttent Net, plays a pivotal role in refining the OD and OC segmentation process. This component acts as the counterpart to the encoder, reconstructing high-resolution feature maps from the abstract representations generated by the encoder. Through a series of up sampling and concatenation operations, the decoder path enables the model to localize and delineate the boundaries of the OD and OC with enhanced precision and accuracy.

5.3.2 Results and Discussion

The performance evaluation of our glaucoma diagnosis model across different factors provides a comprehensive understanding of its capabilities in various clinical contexts. The Area Under the Curve (AUC), sensitivity and specificity metrics offer insights into the model's discriminative power and its ability to accurately classify different severity levels of glaucoma.

In the scenario of a Healthy Fundus, our model demonstrates outstanding performance with an AUC of 99.5%. The remarkable AUC score signifies the model's exceptional capacity to differentiate between healthy and glaucomatous fundi. The sensitivity and specificity metrics, at 98.7% and 99.2% respectively, emphasize the model's precision in accurately detecting healthy fundus cases while reducing false positives.

For the Highly Glaucomatous scenario, our model maintains a robust AUC of 98.9%, indicative of its efficacy in detecting severe glaucomatous conditions. The high sensitivity of 99.5% emphasizes the ability of the model to effectively capture instances that are true positive in highly glaucomatous fundi. The specificity of 97.8% further highlights the ability of the model to reduce the false positives in this challenging scenario.

In the context of Less Glaucomatous fundi, our model achieves an AUC of 97.8%, showcasing its consistent performance across varying severity levels. The sensitivity of 97.2% indicates the model's proficiency in identifying true positive instances in less severe glaucomatous conditions. With a specificity of 98.5%, the model demonstrates its capacity to maintain a high true negative rate, minimizing misclassifications in this scenario.

In scenarios involving Moderate Glaucomatous fundi, our model achieves an AUC of 98.3%, reflecting its robust discriminative power. The sensitivity of 98.6% highlights the model's effectiveness in correctly identifying moderate glaucomatous cases, while the specificity of 98.1% emphasizes its ability to minimize false positives. The performance is summarized in table 5.6.

Table 5.6: Performance of the model for glaucoma diagnosis in different factors detection

Factors	**AUC (%)**	**Sensitivity (%)**	**Specificity (%)**
Healthy Fundus	99.5	98.7	99.2
Highly Glaucomatous	98.9	99.5	97.8
Less Glaucomatous	97.8	97.2	98.5
Moderate Glaucomatous	98.3	98.6	98.1

Table 5.7 offers a comparison between the proposed method and existing ones across various performance indicators.

Table 5.7: Performance comparison of the model with other models

Method		**Accuracy**	**Sensitivity**	**Specificity**	**AUC**
SEDC [68]		0.971	0.952	0.982	0.993
Attention U-net with ResNet-50 [69]		0.957	0.968	0.875	0.98
Inception V3+ResNet50 based U-net Segmentation [70]	ORIGA	0.9117	0.913	0.999	0.90
	REFUGE	0.911	0.937	0.846	0.93
	DRISHTI	0.809	0.833	0.722	0.83
FGSA-HDNN [51]		0.987	0.9964	0.9784	
EH+FCTH [71]		0.804			0.884
Proposed	**REFUGE**		**0.98**	**0.93**	
	ORIGA		**0.99**	**0.96**	
	G1020		**0.97**	**0.95**	

6. Conclusion

Retinal diseases, such as Diabetic Retinopathy (DR), Diabetic Macular Edema (DME), and glaucoma represent a significant threat to both the vision and overall ocular health. Consequently, retinal diseases require early detection and timely intervention to prevent vision loss. The advancements in DL techniques and attention mechanisms presented in this study represent significant milestones in the field of ophthalmic diagnostics. The development of novel frameworks such as the FSSCM, SSCSAC-Net and GlaucoAttent Net showcases substantial progress in automated disease classification and severity assessment for DR, DME and glaucoma.

These innovations leverage TL, FT and self-supervised learning to enhance diagnostic accuracy, reduce dependency on large annotated datasets, and improve the efficiency of disease detection. By integrating advanced preprocessing techniques and attention mechanisms within deep learning architectures, these models achieve superior performance metrics across multiple datasets, setting new benchmarks in clinical reliability and scalability.

The DR disease severity is classified into five classes such as normal, mild, moderate, severe and PDR. The severity detection is acquired through the frame-wise severity scale classification (FSSCM) model which performs in four phases: pre-processing, segmentation, feature extraction and model training. The experiment is carried out by leveraging the pre-trained networks like ResNet101, EfficientNet B3, with the performance validated by computing accuracy and kappa coefficient as 98.9% and 0.98. Similarly, the DME classification encounters challenges due to imbalanced retinal images which are effectively addressed through the proposed model. The scheme employs a weighted cross-entropy function to overcome the class imbalance and the experiment conducted on SSOMDL-Net, CSA-Net and SSCSAC-net yielded enhanced results, particularly for less prevalent classes of DME. Benchmark dataset testing results indicate that the SSCSAC-Net surpasses other techniques across various performance evaluation measures with an accuracy of 98.7%, precision of 98.6% and recall of 98.8%. After that, glaucoma detection and severity classification are performed using GlaucoAttent Net. The severity classification employs an ensemble learning strategy that effectively segments the OC and OD severity. The proposed model demonstrates the enhanced result over the existing schemes. Efficacy metrics showcase outstanding performance: 99.67% accuracy, 99.20% precision, 99.50% recall and a 99.35% F-

score, solidifying the methodology's potential for transformative impact in advancing clinical diagnostics**.**

The impact of these advancements extends beyond technical innovation, promising to revolutionize ophthalmic care by enabling earlier disease detection, more accurate diagnosis, and personalized treatment strategies. By empowering clinicians with AI-driven diagnostic tools that integrate seamlessly into clinical workflows, these models can contribute to improved patient outcomes and enhanced healthcare delivery.

7. Future Scope

However, further improvements are necessary in retinal disease detection including age-related detection such as macular edema, cataract etc., automated retinal classification strategies are essential, and optimization techniques to reduce the model complexity. Moreover, incorporating the advanced techniques like explainable AI and Generative AI schemes will enhances detection of the retinal disease severity more effectively.

As machine learning models become increasingly sophisticated, there is a growing need for explainable AI techniques that can provide transparent explanations for model predictions and recommendations. Integrating explainable AI methods into clinical decision support systems can enhance trust, transparency, and collaboration between healthcare providers and AI systems, ultimately improving patient care outcomes.

Future research in ophthalmic diagnostics could focus on integrating AI models with Electronic Health Records (EHR) for longitudinal patient monitoring, exploring Explainable AI (XAI) techniques to enhance interpretability, and fusing multi-modal imaging data for comprehensive disease assessment. Robustness studies across diverse populations, personalized medicine applications, and real-time clinical decision support systems are also crucial. Collaborative efforts in data sharing and ethical considerations will ensure responsible AI deployment. These advancements aim to optimize diagnostic accuracy, personalize treatment plans, and streamline clinical workflows, ultimately improving patient outcomes and advancing the field of ophthalmology with AI-driven innovations.

8. Publications

Journal

1. Chavan, S., Choubey, N. An automated diabetic retinopathy of severity grade classification using transfer learning and fine-tuning for fundus images. Multimedia Tools Appl 82, 36859–36884 (2023). https://doi.org/10.1007/s11042-023-15135-0
2. Chavan, S., Choubey, N. Self-supervised category selective attention classifier network for diabetic macular edema classification. *Acta Diabetologica* (2024). https://doi.org/10.1007/s00592-024-02257-6

Under Review

1. Glaucoma Detection and Severity Classification based on GlaucoAttent Net framework. submitted to Computers and Electrical Engineering on 02nd April 2024.

Conference Presented

1. Paper entitled “Diagnosis of Eye Disease through Fundus Image Analysis and Machine Learning: Study” presented in ICGETEI-2021 organized by Amity University Rajasthan, Jaipur during 5-6 March 2021.
2. Paper entitled “Comparative analysis of different Machine Learning techniques for diagnosis of eye diseases” presented in ICDTV: 2030 organized by SVKM’s NMIMS, Shirpur campus during 17-18 February, 2024.

9. References

1. Hemanth DJ, Deperlioglu O, Kose U (2020) An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. Neural Computing and Applications. 32(3):707-21.
2. Islam M, Dinh AV, Wahid KA (2017) Automated diabetic retinopathy detection using bag of words approach. Journal of Biomedical Science and Engineering. 10(5):86-96.
3. Qiao L, Zhu Y, Zhou H (2020) Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms. IEEE Access. 8:104292-302.
4. Ishtiaq U, Kareem SA, Abdullah ER, Mujtaba G, Jahangir R, Ghafoor HY (2020) Diabetic retinopathy detection through artificial intelligent techniques: a review and open issues. Multimedia Tools and Applications. 79(21):15209-52.
5. Das S, Kharbanda K, Suchetha M, Raman R, Dhas E (2021) Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy. Biomedical Signal Processing and Control. 68:102600.
6. Costa P, Galdran A, Smailagic A, Campilho A (2018) A weakly-supervised framework for interpretable diabetic retinopathy detection on retinal images. IEEE Access. 6:18747-58.
7. Doshi D, Shenoy A, Sidhpura D, Gharpure P (2016) Diabetic retinopathy detection using deep convolutional neural networks. In2016 International Conference on Computing, Analytics and Security Trends (CAST), IEEE, 261-266.
8. Seewoodhary M (2020) An overview of diabetic retinopathy and other ocular complications of diabetes mellitus. Eye.
9. Lam, C., et al., 2018. Automated detection of diabetic retinopathy using deep learning. AMIA Summits on Translational Science Proceedings, 147-155.
10. Krizhevsky, A., et al., 2012. ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems, 1097-1105.
11. Szegedy, C., et al., 2015. Going deeper with convolutions. In Proc. of IEEE Conference on Computer Vision and Pattern Recognition, 1-9.
12. Li, F., et al. 2019. Automatic Detection of Diabetic Retinopathy in Retinal Fundus Photographs Based on Deep Learning Algorithm. Translational Vision Science & Technology, 8(6):4, 1-13.

13. Iyer SS, Radhakrishnan NS, Roohipourmoallai R, Guerin CM, Maylath JS, Garson N. Chronic ocular small vessel disease: An overview of diabetic retinopathy and its relationship with cardiovascular health. American Heart Journal Plus: Cardiology Research and Practice. 2023 Feb 9:100270.
14. Semeraro F, Morescalchi F, Cancarini A, Russo A, Rezzola S, Costagliola C. Diabetic retinopathy, a vascular and inflammatory disease: therapeutic implications. Diabetes & metabolism. 2019 Dec 1;45(6):517-27.
15. Lanzetta P, Sarao V, Scanlon PH, Barratt J, Porta M, Bandello F, Loewenstein A. Fundamental principles of an effective diabetic retinopathy screening program. Acta diabetologica. 2020 Jul; 57:785-98.
16. Liu X, Ali TK, Singh P, Shah A, McKinney SM, Ruamviboonsuk P, Turner AW, Keane PA, Chotcomwongse P, Nganthavee V, Chia M. Deep learning to detect OCT-derived diabetic macular edema from color retinal photographs: a multicenter validation study. Ophthalmology Retina. 2022 May 1;6(5):398-410.
17. Li X, Hu X, Yu L, Zhu L, Fu CW, Heng PA. CANet: cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading. IEEE transactions on medical imaging. 2019 Nov 6;39(5):1483-93.
18. Lalithadevi B, Krishnaveni S. Detection of diabetic retinopathy and related retinal disorders using fundus images based on deep learning and image processing techniques: A comprehensive review. Concurrency and Computation: Practice and Experience. 2022 Aug 30;34(19):e7032.
19. Sundaram S, Selvamani M, Raju SK, Ramaswamy S, Islam S, Cha JH, Almujally NA, Elaraby A. Diabetic Retinopathy and Diabetic Macular Edema Detection Using Ensemble Based Convolutional Neural Networks. Diagnostics. 2023 Mar 6;13(5):1001.
20. Brito P, Costa J, Gomes N, Costa S, Correia-Pinto J, Silva R. Serological inflammatory factors as biomarkers for anatomic response in diabetic macular edema treated with anti-VEGF. Journal of Diabetes and its Complications. 2018 Jul 1;32(7):643-9.
21. Barros DM, Moura JC, Freire CR, Taleb AC, Valentim RA, Morais PS. Machine learning applied to retinal image processing for glaucoma detection: review and perspective. Biomedical engineering online. 2020 Dec; 19:1-21.

22. Li L, Xu M, Liu H, Li Y, Wang X, Jiang L, Wang Z, Fan X, Wang N. A large-scale database and a CNN model for attention-based glaucoma detection. IEEE transactions on medical imaging. 2019 Jul 8;39(2):413-24.
23. Asaoka R, Hirasawa K, Iwase A, Fujino Y, Murata H, Shoji N, Araie M. Validating the usefulness of the "random forests" classifier to diagnose early glaucoma with optical coherence tomography. American journal of ophthalmology. 2017 Feb 1; 174:95-103.
24. Gheisari S, Shariflou S, Phu J, Kennedy PJ, Agar A, Kalloniatis M, Golzan SM. A combined convolutional and recurrent neural network for enhanced glaucoma detection. Scientific reports. 2021 Jan 21;11(1):1945.
25. Niwas SI, Lin W, Bai X, Kwoh CK, Sng CC, Aquino MC, Chew PT. Reliable feature selection for automated angle closure glaucoma mechanism detection. Journal of medical systems. 2015 Mar; 39:1-0.
26. Sudhan MB, Sinthuja M, Pravinth Raja S, Amutharaj J, Charlyn Pushpa Latha G, Sheeba Rachel S, Anitha T, Rajendran T, Waji YA. Segmentation and classification of glaucoma using U-net with deep learning model. Journal of Healthcare Engineering. 2022 Feb 16;2022.
27. Zago GT, Andreão RV, Dorizzi B, Salles EO (2020) Diabetic retinopathy detection using red lesion localization and convolutional neural networks. Computers in biology and medicine. 116:103537.
28. Shanthi T, Sabeenian RS (2019) Modified Alexnet architecture for classification of diabetic retinopathy images. Computers & Electrical Engineering. 76:56-64.
29. Wu Z, Shi G, Chen Y, Shi F, Chen X, Coatrieux G, Yang J, Luo L, Li S (2020) Coarse-to-fine classification for diabetic retinopathy grading using convolutional neural network. Artificial Intelligence in Medicine. 108:101936.
30. Riaz H, Park J, Choi H, Kim H, Kim J (2020) Deep and densely connected networks for classification of diabetic retinopathy. Diagnostics. 10(1):24.
31. Gayathri S, Gopi VP, Palanisamy P (2020) Automated classification of diabetic retinopathy through reliable feature selection. Physical and Engineering Sciences in Medicine. 43(3):927-45.

32. Gayathri S, Krishna AK, Gopi VP, Palanisamy P (2020) Automated binary and multiclass classification of diabetic retinopathy using Haralick and multiresolution features. IEEE Access. 8:57497-504.
33. Shankar K, Perumal E, Vidhyavathi RM (2020) Deep neural network with moth search optimization algorithm-based detection and classification of diabetic retinopathy images. SN Applied Sciences. 2(4):1-0.
34. Wu T, Liu L, Zhang T, Wu X. Deep learning-based risk classification and auxiliary diagnosis of macular edema. Intelligence-Based Medicine. 2022 Jan 1; 6:100053.
35. Wu J, Zhang Q, Liu M, Xiao Z, Zhang F, Geng L, Liu Y, Wang W. Diabetic macular edema grading based on improved Faster R-CNN and MD-ResNet. signal, image and video processing. 2021 Jun; 15:743-51.
36. Altan G. DeepOCT: An explainable deep learning architecture to analyze macular edema on OCT images. Engineering Science and Technology, an International Journal. 2022 Oct 1; 34:101091.
37. Saini DJ, Sivakami R, Venkatesh R, Raghava CS, Dwarkanath PS, Anwer TM, Smirani LK, Ahammad SH, Pamula U, Hossain MA, Rashed AN. Convolution neural network model for predicting various lesion-based diseases in diabetic macula edema in optical coherence tomography images. Biomedical Signal Processing and Control. 2023 Sep 1; 86:105180.
38. Madadi Y, Abu-Serhan H, Yousefi S. Domain Adaptation-Based deep learning model for forecasting and diagnosis of glaucoma disease. Biomedical Signal Processing and Control. 2024 Jun 1; 92:106061.
39. Yan M, Lin Y, Peng X, Zeng Z. mix DA: mixup domain adaptation for glaucoma detection on fundus images. Neural Computing and Applications. 2023 Jul 24:1-20.
40. Pathan S, Kumar P, Pai RM, Bhandary SV. An automated classification framework for glaucoma detection in fundus images using ensemble of dynamic selection methods. Progress in Artificial Intelligence. 2023 Sep;12(3):287-301.
41. Rasheed HA, Davis T, Morales E, Fei Z, Grassi L, De Gainza A, Nouri-Mahdavi K, Caprioli J. DDLSNet: A novel deep learning-based system for grading funduscopic images for glaucomatous damage. Ophthalmology Science. 2023 Jun 1;3(2):100255.

42. Liu Y, Yip LW, Zheng Y, Wang L. Glaucoma screening using an attention-guided stereo ensemble network. Methods. 2022 Jun 1; 202:14-21.
43. Mrad Y, Elloumi Y, Akil M, Bedoui MH. A fast and accurate method for glaucoma screening from smartphone-captured fundus images. Irbm. 2022 Aug 1;43(4):279-89.
44. Haider A, Arsalan M, Park C, Sultan H, Park KR. Exploring deep feature-blending capabilities to assist glaucoma screening. Applied Soft Computing. 2023 Jan 1; 133:109918.
45. Neto A, Camera J, Oliveira S, Cláudia A, Cunha A. OD and OC segmentations for glaucoma assessment using cup-to-disc ratio. Procedia Computer Science. 2022 Jan 1; 196:485-92.
46. Hervella ÁS, Rouco J, Novo J, Ortega M. End-to-end multi-task learning for simultaneous OD and OC segmentation and glaucoma classification in eye fundus images. Applied Soft Computing. 2022 Feb 1; 116:108347.
47. Vinayaki VD, Kalaiselvi RJ. Multi threshold image segmentation technique using remora optimization algorithm for diabetic retinopathy detection from fundus images. Neural Processing Letters. 2022 Jun;54(3):2363-84.
48. Thanki R. A deep neural network and machine learning approach for retinal fundus image classification. Healthcare Analytics. 2023 Nov 1; 3:100140.
49. Veena HN, Muruganandham A, Kumaran TS. A novel OD and OC segmentation technique to diagnose glaucoma using deep learning convolutional neural network over retinal fundus images. Journal of King Saud University-Computer and Information Sciences. 2022 Sep 1;34(8):6187-98.
50. Feng W, Wang L, Ju L, Zhao X, Wang X, Shi X, Ge Z. Unsupervised domain adaptive fundus image segmentation with category-level regularization. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2022 Sep 16 (pp. 497-506). Cham: Springer Nature Switzerland.
51. Madhumalini M, Devi TM. Detection of Glaucoma from Fundus Images Using Novel Evolutionary-Based Deep Neural Network. Journal of Digital Imaging. 2022 Aug;35(4):1008-22.
52. Agboola HA, Zaccheus JE. Wavelet image scattering based glaucoma detection. BMC Biomedical Engineering. 2023 Mar 2;5(1):1.

53. de Zarzà I, de Curtò J, Calafate CT. Detection of glaucoma using three-stage training with EfficientNet. Intelligent Systems with Applications. 2022 Nov 1; 16:200140.
54. D'Souza G, Siddalingaswamy PC, Pandya MA. AlterNet-K: a small and compact model for the detection of glaucoma. Biomedical Engineering Letters. 2024 Jan;14(1):23-33.
55. Wu J, Fang H, Li F, Fu H, Lin F, Li J, Huang Y, Yu Q, Song S, Xu X, Xu Y. Gamma challenge: glaucoma grading from multi-modality images. Medical Image Analysis. 2023 Dec 1; 90:102938.
56. Gayathri S, Gopi VP, Palanisamy P (2021) Diabetic retinopathy classification based on multipath CNN and machine learning classifiers. Physical and Engineering Sciences in Medicine. 1-5.
57. Gayathri S, Gopi VP, Palanisamy P (2020) A lightweight CNN for Diabetic Retinopathy classification from fundus images. Biomedical Signal Processing and Control. 62:102115.
58. Li YH, Yeh NN, Chen SJ, Chung YC (2019) Computer-assisted diagnosis for diabetic retinopathy based on fundus images using deep convolutional neural network. Mobile Information Systems, 2019.
59. Zeng X, Chen H, Luo Y, Ye W (2019) Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network. IEEE Access. 7:30744-53.
60. Lam C, Yi D, Guo M, Lindsey T (2018) Automated detection of diabetic retinopathy using deep learning. AMIA summits on translational science proceedings. 2018:147.
61. Sayres R, Taly A, Rahimy E, Blumer K, Coz D, Hammel N, Krause J, Narayanaswamy A, Rastegar Z, Wu D, Xu S (2019) Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy. Ophthalmology. 126(4):552-64.
62. Islam MR, Abdulrazak LF, Nahiduzzaman M, Goni MO, Anower MS, Ahsan M, Haider J, Kowalski M. Applying supervised contrastive learning for the detection of diabetic retinopathy and its severity levels from fundus images. Computers in Biology and Medicine. 2022 Jul 1;146:105602.
63. Bhardwaj C, Jain S, Sood M. Deep learning–based diabetic retinopathy severity grading system employing quadrant ensemble model. Journal of Digital Imaging. 2021 Apr;34:440-57.
64. Alahmadi MD. Texture attention network for diabetic retinopathy classification. IEEE Access. 2022 May 23;10:55522-32.

65. Farag MM, Fouad M, Abdel-Hamid AT. Automatic severity classification of diabetic retinopathy based on denseNet and convolutional block attention module. IEEE Access. 2022 Apr 6;10:38299-308.
66. Bilal A, Sun G, Li Y, Mazhar S, Khan AQ. Diabetic retinopathy detection and classification using mixed models for a disease grading database. IEEE Access. 2021 Feb 1;9:23544-53.
67. Al-Antary MT, Arafa Y. Multi-scale attention network for diabetic retinopathy classification. IEEE Access. 2021 Apr 2;9:54190-200.
68. Zhao R, Chen X, Chen Z, Li S. Diagnosing glaucoma on imbalanced data with self-ensemble dual-curriculum learning. Medical Image Analysis. 2022 Jan 1;75:102295.
69. Shyamalee T, Meedeniya D. Glaucoma detection with retinal fundus images using segmentation and classification. Machine Intelligence Research. 2022 Dec;19(6):563-80.
70. Shyamalee T, Meedeniya D, Lim G, Karunarathne M. Automated Tool Support for Glaucoma Identification with Explainability Using Fundus Images. IEEE Access. 2024 Jan 30.
71. Alice K, Deepa N, Devi T, BeenaRani BB, Nagaraju V. Effect of multi filters in glucoma detection using random forest classifier. Measurement: Sensors. 2023 Feb 1;25:100566.